

## CLAIMS

1. A method for producing an audiovisual work, the method comprising the steps of:

providing an audio signal to a speech recognition module;

performing a speech recognition of said audio signal, the speech recognition comprising an extracting of a plurality of basic units of recognized speech and related time codes;

receiving the plurality of basic units of recognized speech and the related time codes from the speech recognition module;

processing the received plurality of basic units to provide synchronization information for a production of said audiovisual work; and

displaying on a user interface said synchronization information.

2. The method as claimed in claim 1, wherein the production comprises post-production audio synchronization, said synchronization information comprises a graphic representation of a sound to be performed at each point in time over a span of time during said audiovisual work, and said interface controls said graphic representation over said span while facilitating synchronized recording of said sound in order to perform post-production.

3. The method as claimed in claim 2, wherein the basic units of recognized speech are phonemes.

4. The method as claimed in claim 2, further comprising the step of converting the basic units of recognized speech received with the time codes into words and words related time codes.

5. The method as claimed in claim 2, further comprising the step of converting the basic units of recognized speech received with the time codes into graphemes and graphemes related time codes, the graphemes being processed to provide synchronization information.

6. The method as claimed in claim 5, further comprising the step of providing a conformed text source, further wherein the synchronization information provided to the user comprises an indication of a temporal location with respect to the audio signal.

7. The method as claimed in claim 5, further comprising the step of providing a script of at least one part of the audio signal, further wherein the synchronization

information provided to the user comprises an indication of a temporal location with respect to the script provided.

8. The method as claimed in claim 5, wherein the displaying on a user interface of said synchronization information, comprises the displaying of the graphemes using a horizontally sizeable font.

9. The method as claimed in claim 5, further comprising the step of detecting a Foley in the audio signal using a Foley detection unit, the detecting comprising the providing of an indication of the Foley and a related Foley time code.

10. The method as claimed in claim 5, further comprising the step of amending at least one part of the audio signal and audio signal related time codes using at least the graphemes and the synchronization information.

11. The method as claimed in claim 4, further comprising the providing of a plurality of words in accordance with the provided audio signal, the providing being performed by an operator.

12. The method as claimed in claim 11, further comprising the step of amending a recognized word in accordance with the plurality of words provided by the operator.

13. The method as claimed in claim 12, further comprising the step of creating a composite signal comprising at least the amended word, a video signal related to the audio source and the audio source.

14. The method as claimed in claim 1, wherein the displaying on a user interface of said synchronization information is used to produce animation.

15. The method as claimed in claim 14, wherein for blocks of continuous spoken word, said synchronization information provides essential visem information for each sequential frame to be drawn by an animator.

16. The method as claimed in claim 15, further comprising the step of providing a storyboard database, further comprising the step of converting the basic units of recognized speech received with the time codes into words and words related time codes, the processing of the plurality of words and the words related time codes providing an indication of a current temporal location of the audio signal with respect to the storyboard.

17. The method as claimed in claim 16, wherein the basic units of recognized speech are phonemes, further comprising the step of providing a plurality of

visems for each of the plurality of words, using a visem database and using the phonemes.

18. The method as claimed in claim 17, further comprising the step of outputting an adjusted voice track comprising the audio signal, at least one part of the storyboard and the plurality of visems.

19. The method as claimed in claim 1, wherein the production comprises adaptation assisting, the adaptation assisting comprises a graphic representation of the plurality of basic units of recognized speech, the related time codes and a plurality of adapted basic units provided by a user, and said interface providing a visual indication of a matching of the plurality of adapted basic units with the plurality of basic speech units, the matching enabling synchronized adaptation of said audio signal.

20. The method as claimed in claim 19, wherein the plurality of adapted basic units is provided by performing a speech recognition of an adapted voice source.

21. The method as claimed in claim 20, wherein the speech recognition of the adapted voice source further provides related adapted time codes, further wherein the step of adapting the audio signal using said synchronization information and the plurality of adapted basic units is performed by attempting to match at least one of the plurality of basic units with at least one of the plurality of adapted basic units using the related time codes and the related adapted time codes.

22. A method for performing closed-captioning of an audio source, the method comprising the steps of:

- providing an audio signal of an audio/video signal to a speech recognition module;

- performing a speech recognition of said audio/video signal, and

- incorporating text of said recognized speech of the audio signal as closed-captioning into a visual or non-visual portion of the audio/video signal in synchronization.

23. The method as claimed in claim 21 further comprising the step of providing an indication of an amount of successful replacement of the plurality of basic units of recognized speech of the audio signal by the plurality of basic units of recognized speech of the adapted audio signal.

24. The method as claimed in claim 23, further comprising the step of providing a minimum amount required of successful replacement of the plurality of basic units of recognized speech of the audio signal by the plurality of basic units of recognized speech of the adapted audio signal, the method further comprising the step of canceling the providing of the at least one replaced plurality of basic units with related replaced time codes if the at least one replaced plurality of basic units is lower than the minimum amount required of successful replacement.
25. The method as claimed in claim 1, wherein the audio signal comprises a plurality of voices originating from a plurality of actors, further comprising the step of assigning each of the plurality of basic units and the related time codes to a related actor of the plurality of actors.
26. The method as claimed in claim 1, wherein the production comprises closed-captioning production of the audio source, said closed-captioning comprises a graphic representation of the recognized plurality of basic units, the method further comprising the incorporating of at least one of the plurality of basic units as closed-captioning in a visual or non-visual portion of the audio/video portion of the audio/video signal in synchronization.
27. The method as claimed in claim 26, further comprising the step of amending at least one part of the plurality of basic units.
28. The method as claimed in claim 1, further comprising the step of converting the basic units of recognized speech received with the time codes into words and words related time codes, further comprising the step of creating a database comprising a word and related basic units.
29. The method as claimed in claim 28, further comprising the step of amending a word of said database, wherein phonemes of the word and the amended word are substantially the same.
30. The method as claimed in claim 1, further comprising the step of converting the basic units of recognized speech received with the time codes into words and words related time codes, further comprising the step of amending at least one word.
31. The method as claimed in claim 30, further comprises the step of providing a visual indication of a word to amend.

32. The method as claimed in claim 1, wherein the audio signal comprises lyrics that are sung, further wherein the production of said audiovisual work comprises a karaoke generation using said audio signal, said karaoke generation comprises a graphic representation of lyrics to be sung at each point in time over a span of time during said audiovisual work using the plurality of basic units of recognized speech provided and related time codes, together with an index representation of a current temporal position with respect to the graphic representation of the lyrics to be sung.

33. The method as claimed in claim 2, further comprising the step of detecting at least one note encoded in the audio signal according to an encoding scheme, further comprising the providing of the detected at least one note on said graphic representation.